AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

FOOD MICROBIOLOGY

Multi-omics Approach Reveals How Yeast Extract Peptides Shape *Streptococcus thermophilus* Metabolism

Lucas Proust,[a,b] Eloi Haudebourg,[a] Alain Sourabié,[b] Martin Pedersen,[c] Iris Besançon,[b] Véronique Monnet,[a] Vincent Juillard[a]

[a]Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, Jouy-en-Josas, France
[b]Procelys by Lesaffre, Maisons-Alfort, France
[c]Sacco S.r.l., Cadorago, Italy

ABSTRACT Peptides present in growth media are essential for nitrogen nutrition and optimal growth of lactic acid bacteria. In addition, according to their amino acid composition, they can also directly or indirectly play regulatory roles and influence global metabolism. This is especially relevant during the propagation phase to produce high cell counts of active lactic acid bacteria used as starters in the dairy industry. In the present work, we aimed at investigating how the respective compositions of two different yeast extracts, with a specific focus on peptide content, influenced *Streptococcus thermophilus* metabolism during growth under pH-controlled conditions. In addition to free amino acid quantification, we used a multi-omics approach (peptidomics, proteomics, and transcriptomics) to identify peptides initially present in the two culture media and to follow *S. thermophilus* gene expression and bacterial protein production during growth. The free amino acid and peptide compositions of the two yeast extracts differed qualitatively and quantitatively. Nevertheless, the two yeast extracts sustained similar levels of growth of *S. thermophilus* and led to equivalent final biomasses. However, transcriptomics and proteomics showed differential gene expression and protein production in several *S. thermophilus* metabolic pathways, especially amino acid, citrate, urease, purine, and pyrimidine metabolisms. The probable role of the regulator CodY is discussed in this context. Moreover, we observed significant differences in the production of regulators and of a quorum sensing regulatory system. The possible roles of yeast extract peptides on the modulation of the quorum sensing system expression are evaluated.

IMPORTANCE Improving the performance and industrial robustness of bacteria used in fermentations and food industry remains a challenge. We showed here that two *Streptococcus thermophilus* fermentations, performed with the same strain in media that differ only by their yeast extract compositions and, more especially, their peptide contents, led to similar growth kinetics and final biomasses, but several genes and proteins were differentially expressed/produced. In other words, subtle variations in peptide composition of the growth medium can finely tune the metabolism status of the starter. Our work, therefore, suggests that acting on growth medium components and especially on their peptide content, we could modulate bacterial metabolism and produce bacteria differently programmed for further purposes. This might have applications for preparing active starter cultures.

KEYWORDS RNA-Seq, CodY, proteomics, quorum sensing, yeast extract

Streptococcus thermophilus is widely used as a starter in dairy fermentation, mainly for yoghurt and cheese production. Industrial production of *S. thermophilus* (and other lactic acid bacteria) largely uses yeast-based nutrients and, particularly, yeast extract (YE) as key nutrient source, in the culture media. YEs are primary raw materials produced from baker's or brewer's yeast through various industrial processes such as

Citation Proust L, Haudebourg E, Sourabié A, Pedersen M, Besançon I, Monnet V, Juillard V. 2020. Multi-omics approach reveals how yeast extract peptides shape *Streptococcus thermophilus* metabolism. Appl Environ Microbiol 86:e01446-20. https://doi.org/10.1128/AEM.01446-20.

Editor Danilo Ercolini, University of Naples Federico II

Address correspondence to Vincent Juillard, vincent.juillard@inrae.fr.

Received 18 June 2020
Accepted 4 August 2020
Accepted manuscript posted online 7 August 2020
Published 28 October 2020

plasmolysis or autolysis and their possible reinforcement by the addition of proteolytic enzymes (1). YE consists of a mixture of nitrogen in various forms (free amino acids, peptides of various lengths, structures, and compositions), carbohydrates, vitamins, nucleotides, and trace elements. In particular, YEs display a high peptide diversity, specifically, with more than 4,000 distinct oligopeptides ranging from 6 to 30 amino acids in length, together with other peptides of shorter length (2). The growth support of lactic acid bacteria by YEs has been attributed for a long time to the presence of large amounts of free amino acids and peptides, among other factors (3). More recently, the role of peptides as a source of amino acids during growth of *Lactococcus lactis* in ^{15}N-labeled yeast hydrolysate containing a large excess of free nonlabeled amino acids has been quantified (4). Both peptides and free amino acids were consumed during growth. The contribution of peptide-bound amino acids to the biomass formation ranged from 30% to 60%, depending on the amino acid, with the remaining part (70% to 40%) originating from free amino acids (4).

Peptides are mainly considered a source of amino acids used for nitrogen nutrition, irrespective of their biochemical features (length, charge, and hydrophobicity), possible structure, and even composition. Nevertheless, they can also be involved in regulatory mechanisms, via direct or indirect actions. Indirect regulatory action of peptides is illustrated by the CodY regulation mechanism. CodY is a pleiotropic transcriptional regulator found in low-GC content Gram-positive bacteria. Its activity is under the control of the intracellular pool of branched-chain amino acids (BCAAs). This pool depends in part on the intake of BCAA-containing peptides further cleaved by the set of intracellular peptidases. As a result, the BCAA-containing peptides indirectly play a regulatory role by providing CodY signals (5, 6). Peptides may also play a direct role in the regulation of physiological traits via the interaction with a dedicated transcriptional regulator (7–9). That is, for instance, the case of the production of streptide, a cyclic peptide whose synthesis by *S. thermophilus* is under the control of a short hydrophobic peptide (EGIIVIVVG) acting as a pheromone via quorum sensing regulation (10–12).

The peptide content of YEs might be subject to intended variations due to process technologies. For instance, the type of enzyme used to degrade yeast proteins will of course guide the peptide content of the yeast extract. The peptide content of YEs might also be subject to unintended variations due to slight batch-to-batch variations of yet unclear origin. In an industrial bioprocess, production instability is a very critical issue, driving end users and YE manufacturers to deploy efforts to ensure consistent quality across batches. However, all these intended and unintended variations might significantly affect the growth rate and growth yield of lactic acid bacteria (13, 14).

The contribution of various technological parameters on the YE broth fermentation yield of *S. thermophilus* was recently investigated using a metabolomics approach (15). The brand of YE clearly affected the medium metabolite dynamics during the fermentation. Nevertheless, it is not clear to which extent these YE brand variations and therefore their different compositions might impact the metabolism of lactic acid bacteria themselves. Our objective was to evaluate how the peptide content of the YE would interfere with the gene expression and protein production of a lactic acid bacterium using omics approaches. To reach that goal, we used an industrial strain of *S. thermophilus* and two YE-based growth media leading to similar fermentation kinetics and yield but produced with distinct processes. We first checked that the nitrogen contents (free amino acids and peptides) were different in the growth media using high-performance liquid chromatography (HPLC) and peptidomics approaches. Then, we evaluated the gene expression and protein production profiles of the strain during fermentation in the two growth media by transcriptomics and proteomics. The results we obtained indicated that media differing only by their yeast extracts, and, more especially, their peptide contents induced differentially regulatory pathways and programmed the strain differently. This, of course, questions the consequences of these differences on the future behavior of the strain.

TABLE 1 Initial yeast-based nutrient growth medium contents of free amino acid and peptides

	Free amino acids (g/100 g)		Peptides (g/100 g)	
Nutrient	Y1	Y2	Y1	Y2
Asp	2.4	2.0	4.0	6.3
Glu	4.3	2.0	6.6	9.7
Asn	1.0	0.8	c[a]	c
Ser	1.2	1.5	0.9	3.3
Gln	ND[b]	0.3	c	c
His	1.4	ND	ND	1.6
Gly	0.7	1.2	2.2	2.6
Arg	1.5	1.7	0.9	1.1
Thr	1.3	1.3	0.7	3.1
Ala	1.9	3.5	1.3	1.7
Tyr	0.7	0.6	0.5	0.9
Met	0.6	0.4	NA[c]	NA
Val	1.9	1.6	2.1	3.1
Trp	0.5	0.4	NA	NA
Phe	2.6	1.3	0.9	1.6
Ile	1.9	1.4	1.2	1.8
Leu	2.8	2.5	1.8	2.7
Lys	1.6	3.8	3.9	3.5
Pro	1.2	0.9	NA	NA
Cys	ND	0.1	NA	NA
Total amino acids	29.5	27.3	27.0	43.0
Positively charged	3.1	5.5	4.8	4.6
Negatively charged	6.7	4.0	10.6	16.0
BCAAs	6.6	5.5	5.1	7.6

[a]c, converted during acidic hydrolysis (Gln into Glu, Asn into Asp).
[b]ND, not detected.
[c]NA, not analyzed.

RESULTS

The two YEs have different amino nitrogen contents. The present work aimed at investigating the growth of *S. thermophilus* in two distinct YE-based growth media and to compare the metabolic responses of the strain. As a first step of analysis, we examined the initial amino acid contents of the two growth media.

The two YEs contained the same total amounts of free amino acids, almost 30 g per 100 g of YE (Table 1). However, their relative compositions were not exactly similar. Y2 contained a lower proportion of acidic amino acids (aspartic and glutamic acids) than Y1 (15% and 23% of the total, respectively) and a higher proportion of positively charged amino acids (lysine and arginine, 20% and 11% of the total, respectively). As an illustration, glutamic acid was the most abundant free amino acid in Y1 (4.3 g/100 g of YE) and lysine the most abundant one in Y2 (3.8 g/100 g). The amount of BCAAs (isoleucine, leucine, and valine) was slightly higher in Y1 than in Y2 (6.6 and 5.5 g/100 g of YE, respectively).

The quantitative peptide content of the YEs was determined by quantifying their constitutive amino acids after acidic hydrolysis (though it is not possible to quantify all the constitutive amino acids of peptides with this method). Y2 appeared to contain 60% more peptides than Y1. Moreover, their relative compositions of peptide-bound amino acids varied. The proportion of positively charged amino acids was lower in Y2 than in Y1 (11 and 18%, respectively) whereas that of acidic residues was similar (39% and 37%, respectively). Note that these amino acid proportions corresponded to a global composition of peptides and did not reflect individual peptide biochemical properties.

The qualitative oligopeptide composition of the two growth media was estimated by two dimensional (2D) liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis as previously reported (2). The number of distinct oligopeptides longer than 6 residues present in Y2 was 35% lower than that of Y1 (850 $\pm$ 87 and 1,333 $\pm$ 169, respectively, means and standard deviations [SDs] from 6 determinations). Y2, there-

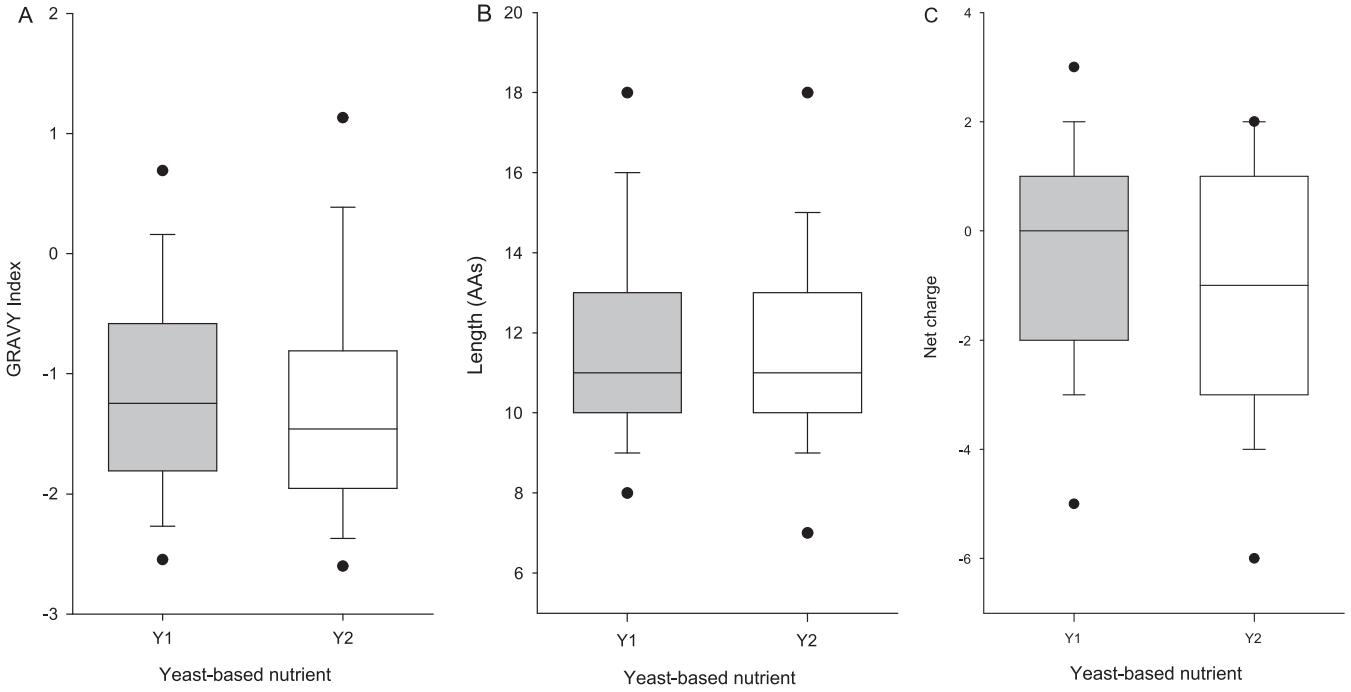

FIG 1 Biochemical properties of peptides identified in YEs. (A) Peptide hydrophobicity. (B) Peptide length. (C) Peptide net charge. ●, outliers as 5th and 95th percentiles.

fore, contained larger amounts of peptides than Y1 but with a lower number of distinct oligopeptide sequences (lower diversity). Moreover, only 30% of the sequences identified in Y2 were also identified in Y1, indicating a significant qualitative difference in terms of oligopeptide compositions between the two YEs. The biochemical features of the two oligopeptide sets have been compared (Fig. 1). The length distributions of the peptides from Y1 and Y2 were comparable. In contrast, peptides from Y2 were slightly statistically more hydrophilic and negatively charged than peptides from Y1.

Bacterial growth and nitrogen consumption dynamics during fermentation. *S. thermophilus* N4L was grown under pH control in Y1 or Y2. Three distinct repetitions were performed. The growth parameters (optical density at 600 nm [OD_{600}] measurement and NaOH consumption) were highly reproducible between repetitions (Fig. 2). Moreover, the growth parameters were similar between the two growth media. Only a slight difference was consistently observed at the end of the exponential phase, which continues a little bit longer in Y2 than in Y1. However, there was no statistical difference between the growth rates of *S. thermophilus* N4L in Y1 and Y2. Moreover, the strain grew to similar final levels in both media, with no significant difference in the final OD values or total amounts of NaOH added.

The time course of amino acid consumption was evaluated during one of the three growth experiments (Fig. 3). The total free amino acid concentration decreased at a higher mean rate in Y2 (2.7 g of amino acid consumed per h) than in Y1 (1.7 g per h). Note that this decrease reflects the balance between free amino acid consumption from the medium by the cells and free amino acid efflux by the cells to the medium and therefore corresponds to a net consumption. The positively charged amino acids were consumed at a higher mean rate in Y2 (0.5 g/h) than in Y1 (0.1 g/h). The mean consumption rates of acidic amino acids were in the same range in the two growth media (0.4 g/h in both cases) as well as that of BCAAs (0.3 and 0.4 g/h, respectively) (data not shown). In contrast, the pool of peptides (regardless of their length) was consumed at a higher mean rate in Y2 (3.4 g/h) than in Y1 (2.4 g/h). Interestingly, the consumption of peptides with a negative net charge was relatively low during fermentation (43% of the initial Y1 pool, 21% of the Y2 initial pool), whereas that of peptides

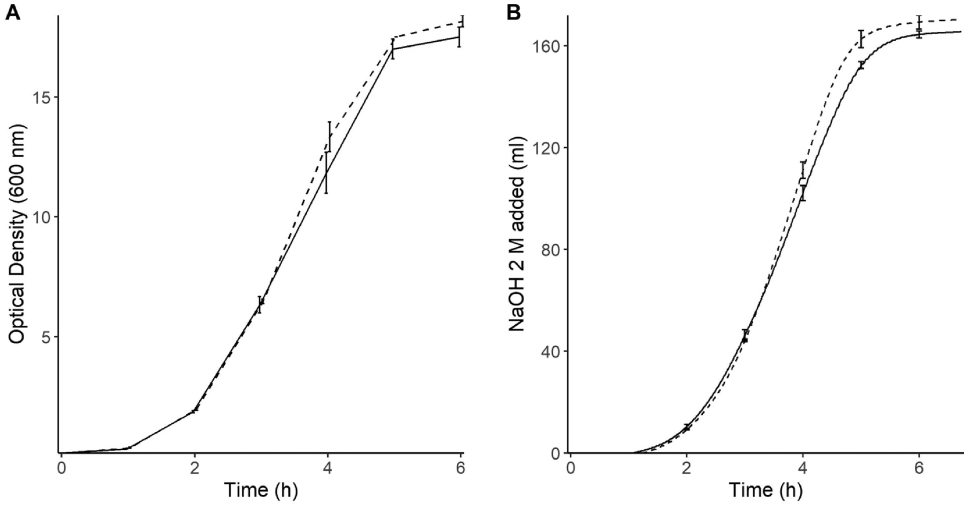

FIG 2 Growth of *S. thermophilus* N4L in YE-containing media performed in 1-liter bioreactors. (A) Cell density evaluated by hourly measurements of optical density at 600 nm. (B) Evolution of the total addition of 2 M NaOH to the culture medium (online monitoring). Means are from 3 independent experiments; error bars, standard deviations; solid line, Y1; dotted line, Y2.

with a positive net charge was much higher (85% and 60% of the initial pools of Y1 and Y2, respectively). Approximately one-half of the peptides containing a BCAA were consumed during growth in Y1 or in Y2 (data not shown). Consequently, despite the similar growth in Y1 and Y2, the consumption profiles of nitrogen compounds by *S. thermophilus* N4L differed significantly between the two media. Growth in Y1 largely relied on the use of free amino acids, which represented a significant part of the nitrogen consumed, whereas the consumption of peptides in Y2 was more significant than in Y1. Overall, the strain consumed a larger amount of total nitrogen (free amino acids and peptides) in Y2 than in Y1 due to a higher global consumption rate (ca. 6 vs. 4 g/h, respectively). In contrast, the growth yields of N4L in the two media were comparable (Fig. 2). It therefore questions the fate of this excess of consumed nitrogen, especially peptides.

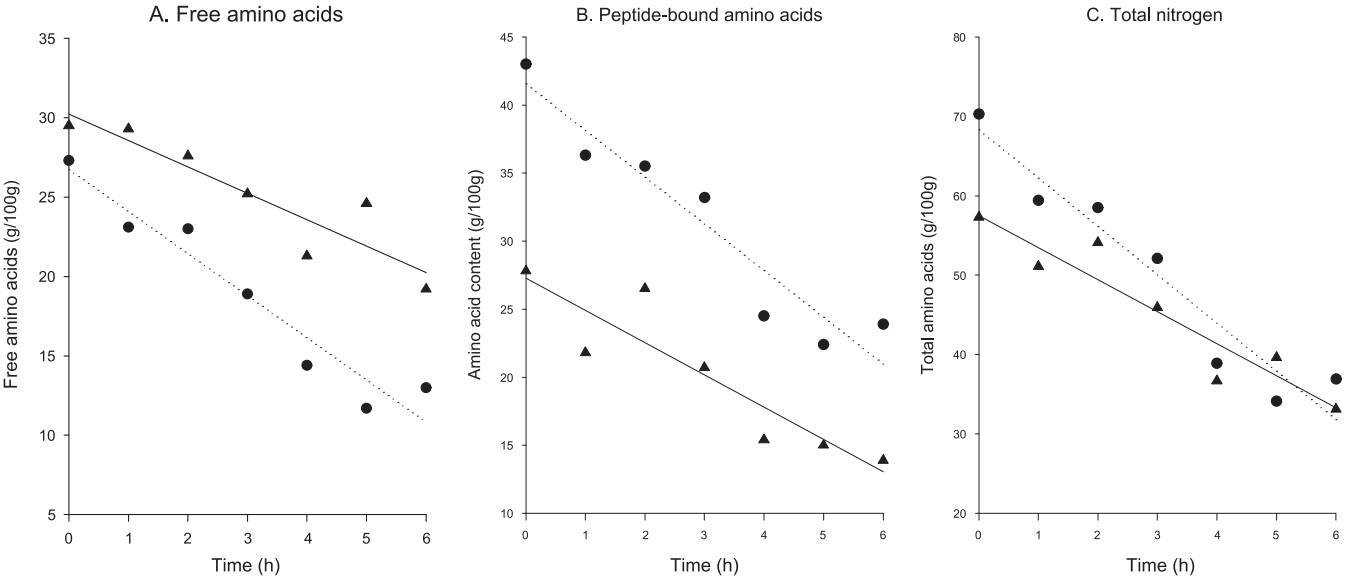

FIG 3 Time course of nitrogen consumption during growth of *S. thermophilus* N4L in Y1 (▲) and Y2 (●). (A) Free amino acids. (B) Peptides. (C) Total amino nitrogen (free amino acids plus peptides).

TABLE 2 Number of identified proteins according to their cellular localization prediction (LocateP)

Cellular localization	No. of proteins in:		
	CP samples	EP samples	Both CP and EP samples
Cytoplasmic	538	575	478
Envelope[a]	40	179	39

[a]Includes membrane-embedded, membrane-anchored, cell wall-anchored, and secreted proteins and lipoproteins.

Overview of transcriptome analysis. Three replicate cultures of *S. thermophilus* N4L were performed for each YE growth medium, Y1 and Y2. Samples were collected after 3, 4, and 5 h of growth, corresponding to a window enclosing the exponential and early stationary growth of the strain (Fig. 2), and used for both transcriptomics and proteomics.

As the presence of short genes is usually underestimated by conventional annotation tools, the *S. thermophilus* N4L genome was processed using the BactgeneSHOW program (16, 17). This program is based on a Hidden Markov Model that uses intrinsic sequence information (presence of ribosome binding sites [RBSs] and nucleotide composition) to predict short putative coding sequences (shCDS). Six hundred seventy-nine shCDS (of 45 to 180 nucleotides [nt]) that were not present in the published annotation of N4L with 2,209 genes were predicted (see Table S1 in the supplemental material).

The construction of two-stranded libraries of regular and short transcripts resulted *in fine* to the generation of averages of 28 million and 23 million of cleaned reads per sample, respectively (see Table S2). A large proportion of the regular transcripts (89%, corresponding to approximately 24 million) was unambiguously assigned to regular genes, whereas this proportion was substantially lower for short reads (1.5 million, i.e., 7% of the cleaned short transcripts). Abundance data from the two libraries were subsequently pooled. The final abundance data set included 1,119 genes (39% of the total number of genes) with a normalized abundance value inferior to 100 reads in any sample. These genes were considered not or poorly expressed and were not analyzed further (see Table S3). The other 1,769 genes (61% of the total) were considered significantly expressed genes. Finally, expressed genes represented 77% and 10% of the annotated and short predicted genes, respectively.

Overview of proteome analysis. The proteome analysis was carried out for the same sampling points as the transcriptomics. According to the protein preparation protocol used, two different protein pools were considered: one contained largely cytoplasmic proteins (CP) and the other was enriched in envelope-bound proteins (EP). Totals of 578 and 754 proteins were identified and fully quantified in the CP and EP samples, respectively, with a large overlap of identification between the two fractions (Table 2). This overlap mainly concerned cytoplasmic proteins that were also detected in the EP fraction, indicating that the EP fraction is only somewhat enriched in envelope proteins rather than containing mainly envelope proteins.

In total, 815 unique proteins were identified in the proteomic analyses, representing ca. 50% of the total encoded proteins. For further analysis, results from CP and EP fractions were pooled. Redundancy of identification was treated as follows: for proteins predicted as being cytoplasmically located, data from the CP fraction were used; for noncytoplasmic proteins, data from EP fraction were considered.

Combined overview of transcriptome and proteome analyses. Comparison of transcriptome and proteome data indicated that the transcriptomics coverage was slightly higher than that of proteomics (detection of 61% of the genes, compared to 50% of the proteins). According to the literature, these values were in the expected range of coverage. However, they enclose all the sampling points, regardless of the fermentation stage or the YE. To have a first idea of the time course of gene expression and protein production as well as the influence of YE, principal-

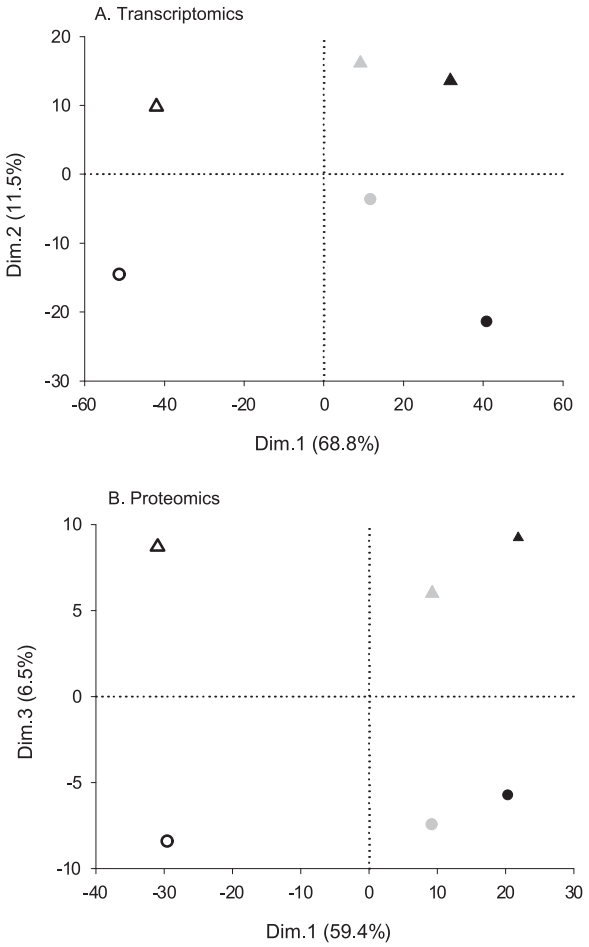

FIG 4 PCAs based on transcript (A) and protein (B) abundances estimated by read counts and extracted-ion chromatography (XIC), respectively. Sampling time points are represented in black (3 h), gray (4 h), and open symbols (5 h). Y1 is shown as triangles, Y2 as circles.

component analyses (PCAs) based on transcriptome and proteome abundance data were calculated (Fig. 4). For both cases, the first component systematically separated samples according to their sampling time point and summarized approximately 60% of total variance. Nevertheless, a separation according to the YE-based growth medium was also observed in the second dimension for transcriptomic data, in the third one for proteomics, accounting for 11.5% and 6.5% of total variance, respectively. These results show that the YE composition also had an impact on the physiological response of N4L, albeit it was more moderate than that of the stage of growth.

Differential impacts of YEs on transcript abundances. The total numbers of differentially expressed genes (DEGs) and differentially produced proteins (DPPs), regardless of the fold change, are given in Table 3. At the transcriptomic level, among the 1,769 genes considered expressed during growth, 969 were statistically differentially expressed between the two growth media, with an adjusted *P* value threshold of 0.05 (Table S3). These genes included 44 unannotated shCDS predicted by BactgeneSHOW. This means that 55% of the genes considered expressed during growth varied differently with Y1 compared to that with Y2. Nevertheless, genes mostly varied as a function of growth time according to PCA (Fig. 4). It therefore suggests that the variations due to the growth medium were of moderate intensity (low fold change). As a matter of fact, arbitrary exclusion of the genes having a fold change (FC) ranging between 0.5 and 2.0 reduced the number of differentially expressed genes from 969 to 96 designated substantially differentially expressed genes (sDEGs). These sDEGs included 38 shCDS and 58 regular genes (Table 4).

TABLE 3 Number of differentially expressed genes and differentially synthesized proteins as a function of growth time[a]

	No. (SD) of genes[b]			No. (SD) of proteins[c]		
Expression/production	**3 h**	**4 h**	**5 h**	**3 h**	**4 h**	**5 h**
Up	335 (9)	150 (3)	227 (8)	43	57	91
Down	307 (16)	131 (18)	193 (13)	37	39	47
Total	642 (25)	281 (21)	420 (21)	80	96	138

[a]Differences were selected at an adjusted *P* value threshold of 0.05, regardless of the fold change.
[b]Only genes showing more than 100 normalized counts in any sample were considered. The number (SD) of unique genes was 969 (44).
[c]Combination of CP- and EP-deriving results. The number of unique proteins was 239.

These regular sDEGs were mainly involved in 4 general metabolism pathways. One of these parts was amino acid metabolism, notably, arginine biosynthesis (4 genes), histidine biosynthesis (3 genes), and glutamine transport (4 genes). In the cases of Gln transport (*STN4L_02239* to *-02242*) and Arg biosynthesis (*STN4L_00969*, *-01129*, *-01963*, and *-01964*), the expression level was higher in Y2 than in Y1 in the exponential growth phase (t3; first sampling point) but higher in Y1 than in Y2 at the end of the exponential growth (t5; last sampling point). Genes involved in His biosynthesis (*STN4L_00335*, *-00337*, and *-00340*) were systematically more highly expressed in Y2 than in Y1, regardless of the growth phase. Moreover, 4 other genes of the His biosynthesis cluster, *STN4L_00336*, *-00338*, *-00339*, and *00343*, were also consistently overexpressed in Y2, albeit with a lower fold change (mean FC of 1.6). It therefore reinforced the observed trend of a strong induction of His biosynthesis in Y2. A gene enrichment analysis indicated that amino acid biosynthesis was the most significantly (adjusted *P* value 0.05) and the most intensively overrepresented pathway, especially histidine and arginine metabolisms (Table 5).

Other metabolism pathways concerned nucleotide biosynthesis (7 genes), citrate metabolism (3 genes), and regulatory genes (3 genes). In particular, the time course of *codY* (*STN4L_02180*) expression was different in the two cultures. The expression level was stable during growth in Y1, whereas it increased by a factor of 3 during growth in Y2 so that the *codY* gene was detected as substantially differentially expressed. Despite the fact that Y2 is reported to contain a larger amount of nucleotides than Y1 (Procelys, personal communication), the expression of the genes of purine metabolism was surprisingly higher in Y2 than in Y1 during the exponential growth phase. However, their expression levels were similar in the two media at the end of growth. In contrast, the genes involved in citrate metabolism were overexpressed in Y1. In addition to these sDEGs, 10 genes of the urease cluster (*STN4L_1336* to *-1345*) were also consistently overexpressed in Y2, even if the mean FC was only 1.4.

Analysis of differentially expressed short genes shed light on four short genes of potential interest (Table 6). *STN4L_short_044* had completely different expression profiles in the two growth media, with an absence of significant expression in Y2 and time-induced expression in Y1 (Fig. 5). *STN4L_short_044* encodes a short hydrophobic pheromone controlling streptide production through a quorum sensing mechanism (10, 11). In agreement with the depicted scheme of streptide production, the four known target genes of STN4L_short_044, namely, *STN4L_short_044* itself (positive regulation loop), *STN4L_00157* (streptide), *STN4L_00158* (streptide maturase), and *STN4L_00159* (streptide exporter) were not induced in Y2, whereas their expression was clearly induced in Y1. In particular, *STN4L_00157* was the strongest differentially expressed gene of the strain, being approximately 60-fold more expressed in Y1 than in Y2 at the end of growth. The expression profile of *STN4L_short_415* was completely different from that of *short_044*, with a significant expression only in Y2, expression that decreased with incubation time. This was also the case for *short_338*. In contrast, expression of *STN4L_short_375* increased as a function of growth time in both YEs, expression intensity being higher in Y1 than Y2. In summary, these short genes did not have the same time-course expression and were not all expressed in the same YE-containing media. The presence of genes encoding transcriptional regulators (*ylbF* and *yebC*) and a response regulator (*rr01*, homolog to *covR*) upstream of

TABLE 4 *S. thermophilus* N4L genes significantly differentially expressed during growth in two yeast extracts

Locus	Annotation	Pathway	FC[a] (Y2/Y1)			log_2 FC		
			t3	t4	t5	t3	t4	t5
STN4L_00055	Amino acid ABC transporter (substrate binding protein)	Amino acid metabolism	1.69	0.46	0.58	0.76	−1.11	−0.78
STN4L_00056	Acetate kinase	Purine metabolism	1.41	0.42	0.40	0.50	−1.25	−1.32
STN4L_00092	Protein of unknown function		0.70	0.57	0.47	−0.52	−0.82	−1.09
STN4L_00094	Oxalate:formate antiporter		2.29	1.09	1.12	1.19	0.12	0.16
STN4L_00157	Protein of unknown function	Streptide metabolism	1.35	0.11	0.02	0.44	−3.20	−5.97
STN4L_00158	KXXXW cyclic peptide radical SAM maturase	Streptide metabolism	1.24	0.33	0.16	0.31	−1.61	−2.66
STN4L_00159	Uncharacterized protein	Streptide metabolism	1.43	0.37	0.23	0.52	−1.45	−2.15
STN4L_00160	Coenzyme PQQ synthesis protein	Coenzyme metabolism	1.43	0.29	0.18	0.52	−1.79	−2.47
STN4L_00162	Uncharacterized protein		1.83	0.41	0.24	0.87	−1.28	−2.05
STN4L_00163	Protein of unknown function		1.15	0.38	0.19	0.20	−1.40	−2.42
STN4L_00193	Uncharacterized protein		2.84	1.92	1.16	1.51	0.94	0.21
STN4L_00208	Protein of unknown function		0.68	0.64	2.03	−0.55	−0.63	1.02
STN4L_00212	Hypothetical protein		0.77	0.40	0.20	−0.38	−1.31	−2.31
STN4L_00220	Uncharacterized protein		1.07	1.56	2.20	0.10	0.64	1.13
STN4L_00223	Uncharacterized protein		1.43	2.04	1.99	0.51	1.03	0.99
STN4L_00290	Aconitate hydratase	Citrate metabolism	0.26	0.66	0.74	−1.93	−0.60	−0.43
STN4L_00291	Citrate synthase	Citrate metabolism	0.27	0.68	0.87	−1.89	−0.55	−0.20
STN4L_00292	Isocitrate dehydrogenase (NADP)	Citrate metabolism	0.28	0.70	0.81	−1.85	−0.52	−0.30
STN4L_00335	Histidinol-phosphate aminotransferase	Amino acid metabolism (His)	1.19	1.66	2.10	0.25	0.73	1.07
STN4L_00337	ATP phosphoribosyltransferase	Amino acid metabolism (His)	1.25	1.84	2.16	0.32	0.88	1.11
STN4L_00340	Imidazole glycerol phosphate synthase subunit HisH	Amino acid metabolism (His)	1.19	1.64	2.00	0.25	0.71	1.00
STN4L_00439	Uncharacterized protein		2.14	1.32	1.18	1.10	0.40	0.24
STN4L_00612	Protein of unknown function		0.53	0.67	0.42	−0.92	−0.57	−1.25
STN4L_00652	Acetoin dehydrogenase	Acetoin metabolism	0.82	0.47	1.07	−0.29	−1.09	0.10
STN4L_00714	Uncharacterized protein		0.21	0.48	2.80	−2.26	−1.06	1.48
STN4L_00715	Protein of unknown function		0.39	0.47	0.59	−1.37	−1.08	−0.77
STN4L_00866	Protein of unknown function		0.34	0.80	0.70	−1.57	−0.32	−0.51
STN4L_00906	Hypothetical protein		0.81	1.27	2.15	−0.30	0.35	1.11
STN4L_00969	Ornithine carbamoyltransferase	Amino acid metabolism (Arg)	1.55	0.47	0.54	0.63	−1.08	−0.88
STN4L_01004	Uncharacterized protein		1.42	0.94	2.12	0.51	−0.10	1.09
STN4L_01063	Carbamoyl-phosphate synthase (glutamine hydrolyzing)	Pyrimidine metabolism	0.37	0.91	0.84	−1.45	−0.14	−0.25
STN4L_01064	Carbamoyl-phosphate synthase small chain	Pyrimidine metabolism	0.48	0.96	0.77	−1.05	−0.06	−0.38
STN4L_01126	Uncharacterized protein		1.72	1.93	2.28	0.78	0.95	1.19
STN4L_01129	Arginine biosynthesis bifunctional protein ArgJ	Amino acid metabolism (Arg)	1.62	0.51	0.49	0.70	−0.96	−1.03
STN4L_01175	Transcriptional regulator	Regulator	0.42	0.82	0.95	−1.27	−0.28	−0.08
STN4L_01277	Hypothetical protein		4.84	2.79	1.73	2.27	1.48	0.79
STN4L_01422	Uncharacterized protein		2.18	1.55	1.42	1.12	0.64	0.51
STN4L_01581	ISSth1, transposase (Orf2), IS*3* family		1.16	2.05	2.19	0.21	1.03	1.13
STN4L_01647	Bifunctional purine biosynthesis protein	Purine metabolism	2.11	0.99	0.97	1.08	−0.01	−0.05
STN4L_01648	Phosphoribosylglycinamide (GAR) formyltransferase	Purine metabolism	2.27	0.97	0.91	1.18	−0.05	−0.13
STN4L_01649	Phosphoribosylformylglycinamidine cyclo-ligase	Purine metabolism	2.55	0.98	0.87	1.35	−0.03	−0.21
STN4L_01650	Amidophosphoribosyltransferase	Purine metabolism	2.50	1.08	0.99	1.32	0.12	−0.01
STN4L_01866	A0A249DMB6 (uncharacterized protein)		2.02	1.05	0.71	1.01	0.07	−0.50
STN4L_01963	Argininosuccinate synthase	Amino acid metabolism (Arg)	1.38	0.41	0.52	0.46	−1.29	−0.93
STN4L_01964	Argininosuccinate lyase	Amino acid metabolism (Arg)	1.80	0.48	0.45	0.85	−1.06	−1.17
STN4L_01998	Transposase (IS*1193*)		1.08	2.30	1.57	0.11	1.20	0.65
STN4L_02026	Uncharacterized protein		0.59	0.42	0.29	−0.76	−1.27	−1.78
STN4L_02035	Protein of unknown function		0.63	0.64	3.23	−0.67	−0.65	1.69
STN4L_02079	Response regulator of the LytR/AlgR family protein	Regulator	0.48	0.93	0.88	−1.07	−0.11	−0.18
STN4L_02180	GTP-sensing transcriptional pleiotropic repressor CodY	Regulator	0.77	1.17	2.07	−0.38	0.22	1.05
STN4L_02182	Lipoprotein		1.57	1.15	2.36	0.65	0.20	1.24
STN4L_02206	Uncharacterized protein		0.40	0.86	1.60	−1.33	−0.22	0.67
STN4L_02239	Amino acid ABC transporter permease	Amino acid metabolism (Gln)	2.87	0.41	0.55	1.52	−1.30	−0.87
STN4L_02240	Amino acid ABC transporter permease	Amino acid metabolism (Gln)	3.45	0.42	0.59	1.79	−1.25	−0.75
STN4L_02241	Glutamine ABC transporter, ATP-binding protein GlnQ	Amino acid metabolism (Gln)	3.47	0.45	0.78	1.80	−1.15	−0.36
STN4L_02242	Amino acid ABC transporter amino acid-binding protein	Amino acid metabolism (Gln)	2.56	0.44	0.56	1.36	−1.19	−0.84
STN4L_02252	Uncharacterized protein		0.89	0.24	0.04	−0.17	−2.06	−4.64
STN4L_02254	Uncharacterized protein		1.23	0.24	0.05	0.30	−2.04	−4.47

[a]FC, fold change.

STN4L_short_375, *-415*, and *-338*, respectively, questions a possible interaction between these encoded short peptides and their adjacently located encoded regulators.

Differential impacts of YEs on protein abundances. At the proteomic level, 184 proteins had an abundance variation between the two YE-based growth media of at

TABLE 5 KEGG enrichment analyses

		Transcriptomics						Proteomics					
		t3		t4		t5		t3		t4		t5	
KEGG no.	Description	*P* value (adjusted)	Gene count	*P* value (adjusted)	Gene count	*P* value (adjusted)	Gene count	*P* value (adjusted)	Protein count	*P* value (adjusted)	Protein count	*P* value (adjusted)	Protein count
ko01230	Biosynthesis of amino acids	3.31E−34	48	2.73E−31	33	8.13E−23	31	3.74E−05	7	5.80E−03	5	5.69E−06	10
ko03010	Ribosome	2.11E−13	23	2.54E−05	9	8.77E−03	7			4.91E−06	7	4.42E−16	16
ko00230	Purine metabolism	4.74E−10	23	2.66E−02	6	5.00E−05	12	1.45E−07	9	4.49E−04	6		
ko00220	Arginine biosynthesis	4.06E−07	11	3.13E−07	8	5.51E−06	8						
ko00250	Alanine, aspartate, and glutamate metabolism	2.08E−06	11			8.77E−03	5	4.83E−02	2				
ko00270	Cysteine and methionine metabolism	5.12E−06	13	1.00E−02	5	4.34E−02	5					2.39E−02	4
ko01210	2-Oxocarboxylic acid metabolism	7.86E−06	11	1.27E−09	11	5.51E−06	9	3.74E−05	5	1.20E−02	3	8.75E−03	4
ko00240	Pyrimidine metabolism	4.65E−05	11			2.37E−05	9						
ko01200	Carbon metabolism	6.22E−05	21	2.92E−03	10			1.24E−02	5				
ko03440	Homologous recombination	1.14E−04	9					2.92E−04	4			4.10E−02	3
ko03430	Mismatch repair	2.10E−04	7					6.10E−05	4			1.30E−03	4
ko00521	Streptomycin biosynthesis	2.40E−04	5										
ko00340	Histidine metabolism	2.40E−04	7	1.64E−09	9	8.79E−08	9						
ko02024	Quorum sensing	2.82E−04	17	9.30E−03	8								
ko00061	Fatty acid biosynthesis	7.07E−04	6			4.34E−02	3						
ko03060	Protein export	7.07E−04	6	1.68E−02	3								
ko00260	Glycine, serine, and threonine metabolism	8.68E−04	9										
ko00300	Lysine biosynthesis	1.43E−03	6			1.14E−04	6						
ko03420	Nucleotide excision repair	1.72E−03	6										
ko00550	Peptidoglycan biosynthesis	2.07E−03	6			1.62E−06	8						
ko00290	Valine, leucine, and isoleucine biosynthesis	2.44E−03	4	1.31E−04	4	5.08E−04	4	1.44E−04	3	1.20E−02	2		
ko00670	One-carbon pool by folate	3.16E−03	5			4.06E−02	3	1.74E−02	2				
ko00710	Carbon fixation in photosynthetic organisms	3.16E−03	5										
ko03030	DNA replication	5.17E−03	6					3.36E−03	3			2.81E−02	3
ko03070	Bacterial secretion system	1.47E−02	6										
ko00620	Pyruvate metabolism	1.76E−02	7			3.49E−02	5						
ko00900	Terpenoid backbone biosynthesis	2.32E−02	5			2.63E−02	4						
ko01212	Fatty acid metabolism	2.46E−02	6										
ko00770	Pantothenate and CoA biosynthesis	2.53E−02	4	1.72E−03	4	7.44E−03	4	1.76E−02	2	3.23E−02	2		
ko01502	Vancomycin resistance	3.00E−02	3										
ko00970	Aminoacyl-tRNA biosynthesis	3.28E−02	5			1.42E−06	9						
ko00020	Citrate cycle (TCA cycle)			7.05E−05	6			3.65E−02	2			2.90E−03	4
ko00660	C_5-branched dibasic acid metabolism			6.59E−04	4								
ko00400	Phenylalanine, tyrosine and tryptophan biosynthesis			1.80E−03	5	4.34E−02	4						
ko00010	Glycolysis/gluconeogenesis			7.74E−03	5								
ko00791	Atrazine degradation			2.89E−02	2								
ko01502	Vancomycin resistance					1.10E−02	3						
ko00983	Drug metabolism, other enzymes					3.28E−02	3						
ko00720	Carbon fixation pathways in prokaryotes							1.24E−02	3				

TABLE 6 Relative expression values of some short and adjacent genes

Locus	Strand	Name	Product	Avg relative expression ± SD					
				Y1			Y2		
				3 h	4 h	5 h	3 h	4 h	5 h
STN4L_short_044	Reverse	*shp*	MKKQILLTLLLVVFEGIIVIVVG[a]	4 ± 2	41 ± 19	173 ± 89	3 ± 2	3 ± 1	6 ± 3
STN4L_00156	Forward	*rgg*	HTH-type transcriptional regulator rgg	1,616 ± 224	1,633 ± 112	2,205 ± 357	1,529 ± 168	1,769 ± 130	3,064 ± 505
STN4L_00157	Forward	*strA*	Streptide	759 ± 92	18,033 ± 10,392	70,998 ± 39,432	1,026 ± 191	1,957 ± 265	1,161 ± 902
STN4L_00158	Forward	*strB*	KXXXW cyclic peptide radical SAM maturase	1,201 ± 246	3,916 ± 1778	4,356 ± 1,691	1,492 ± 157	1,279 ± 93	689 ± 74
STN4L_00159	Forward	*strC*	Dedicated ABC exporter	449 ± 101	1,425 ± 670	1,309 ± 559	644 ± 88	521 ± 47	295 ± 86
STN4L_01261	Reverse		YlbG	185 ± 53	175 ± 20	115 ± 11	227 ± 11	219 ± 15	91 ± 20
STN4L_01262	Reverse		YlbF family regulator	7,490 ± 1,147	8,541 ± 854	3,890 ± 393	9,662 ± 1,003	7,746 ± 310	2,586 ± 255
STN4L_short_375	Reverse		MLIGLYKNRFIWLFVL	110 ± 22	167 ± 22	440 ± 16	15 ± 3	77 ± 5	331 ± 191
STN4L_01310	Reverse	*hk01*	Histidine kinase	4,637 ± 268	4,062 ± 238	1,237 ± 204	5,654 ± 158	4,326 ± 213	1,027 ± 155
STN4L_01311	Reverse	*rr01*	Response regulator (*covR* homolog)	18,050 ± 249	13,062 ± 597	4,890 ± 554	16,431 ± 223	13,429 ± 442	4,207 ± 585
STN4L_01312	Reverse		DNA-binding protein	8,029 ±1,037	6,674 ± 463	2,386 ± 400	10,004 ± 787	7,032 ± 365	1,848 ± 291
STN4L_short_388	Reverse		MLIKSKIIITRFNEK	47 ± 12	50 ± 11	35 ± 10	122 ± 32	92 ± 27	31 ± 15
STN4L_01440	Reverse	*yebC*	Transcriptional regulator	34,857 ± 233	32,882 ± 914	13,778 ± 1,154	38,824 ± 2,029	34,648 ±1,325	15,388 ± 700
STN4L_short_415	Reverse		VNLAGSAIFKKEKNKSEML	25 ± 5	45 ± 20	28 ± 9	381 ± 49	437 ± 130	196 ± 120

[a]Sequence for the mature form of the pheromone is underlined.

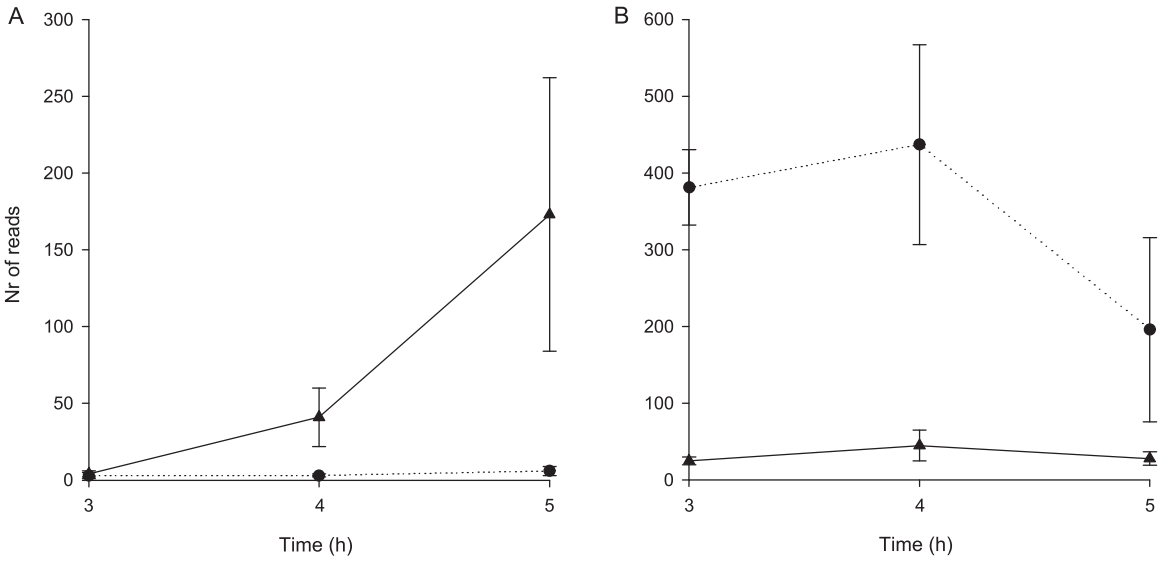

FIG 5 Time course expression of the pheromone activating streptide production (A) and the short gene 415 (B). Y1 is shown as triangles, Y2 as circles.

least 50% in at least one sampling point (see Table S4), regardless of the adjusted *P* value. The abundance of 99 of them was statistically different (*P* adjusted $\leq$ 0.05) between Y1 and Y2, and the most significant changes are reported in Table 7. Proteins involved in citrate metabolism were largely overproduced in Y1, with FC ranging from 4.1 to 6.8 for citrate synthase (GltA), higher than 3.6 for isocitrate dehydrogenase (icd) at two sampling points (t3 and t5) and higher than 2.5 for aconitate hydratase (CitB) at t5. These observations were therefore in agreement with the results of gene expression analysis. The stress resistance proteins AhpD and UmuD (alkyl hydroperoxide reductase and UV resistance protein, respectively) were also systematically overproduced in Y1, with fold changes higher than 1.6. In contrast, proteins involved in purine metabolism (PurD, -H, -M, -F, and -C) were systematically overproduced in Y2, with FCs frequently higher than 2. These results were again consistent with the transcriptomics data. The sensor histidine kinase HK02 was also systematically more abundantly produced in Y2 than in Y1. This sensor belongs to a two-component system that seems to be specific to *S. thermophilus*, whose function has not been deciphered yet (18). The protein enrichment analysis mainly highlighted amino acid, purine, and DNA metabolism pathways (Table 5). As expected, two of these pathways were already underlined during transcriptomic analysis.

Comparison of DEGs and DPPs yielded 9 genes/proteins that were present in both analyses. Three of them were involved in citrate metabolism (*citB*, *gltA*, and *icd*) and 4 in purine metabolism (*purH*, *purM*, *purF*, and *carA*, also involved in amino acid metabolism). The two remaining genes were involved in regulatory processes (*lytR*) and amino acid transport and metabolism (*glnH_1*). Expanding the comparison of the genes and proteins that were differentially expressed/produced (at least 50% of variation) regardless of their adjusted *P* value resulted in 19 supplementary genes/proteins. Five of them were involved in amino acid metabolism (*hisD*, *argG*, *cysM1*, *dagA*, and *STN4L*_00516), 3 in purine metabolism (*purC*, *purK*, and *purE*), and 3 others in DNA metabolism (*infA*, *rpmJ*, and *addA*).

DISCUSSION

Bacterial gene expression and protein production have been analyzed during the growth of *S. thermophilus* N4L in media with two different YEs. Analysis of the nitrogen content (free amino acids and peptides) revealed some differences between the two YEs in terms of concentrations and compositions. This was most probably due to differences in the technological processes used for the two YE productions. YE pro-

TABLE 7 Main *S. thermophilus* N4L proteins significantly differentially produced during growth in the two yeast extracts

Locus	Annotation	Pathway	FC[a] (Y2/Y1)			log_2 FC		
			t3	t4	t5	t3	t4	t5
STN4L_00048	Glutamine ABC uptake transporter membrane-spanning protein	Amino acid metabolism (Gln)		0.5719	0.4817		−0.81	−1.05
STN4L_00290	Aconitate hydratase	Citrate metabolism			0.3915			−1.35
STN4L_00291	Citrate synthase	Citrate metabolism	0.2456	0.2152	0.1463	−2.03	−2.22	−2.77
STN4L_00292	Isocitrate dehydrogenase (NADP)	Citrate metabolism	0.2338		0.2774	−2.10		−1.85
STN4L_00338	Histidinol dehydrogenase	Amino acid metabolism (His)	1.5981	1.8195	2.0292	0.68	0.86	1.02
STN4L_00366	Uncharacterized protein			2.1719			1.12	
STN4L_00379	DNA-directed DNA polymerase	DNA replication	0.5883		0.4501	−0.77		−1.15
STN4L_00402	Chloride channel protein EriC	Ion transport			0.3474			−1.53
STN4L_00418	Competence associated membrane nuclease	DNA metabolism		1.8000	2.0962		0.85	1.07
STN4L_00503	Uncharacterized protein		0.5943	0.6433	0.6434	−0.75	−0.64	−0.64
STN4L_00548	Lipoate-protein ligase	Protein posttransductional modification	0.6407		0.5994	−0.64		−0.74
STN4L_00555	Phosphate ABC uptake transporter membrane-spanning protein	Ion transport	1.5968	1.7588	1.8647	0.68	0.81	0.90
STN4L_00574	Sodium/alanine glycine symporter	Amino acid metabolism		1.6310	1.6410		0.71	0.71
STN4L_00668	UV resistance protein UmuD	Stress resistance	0.5247	0.6070	0.4962	−0.93	−0.72	−1.01
STN4L_00673	Transcriptional regulator	Regulation process		0.6151	0.5800		−0.70	−0.79
STN4L_00698	Zinc ABC transporter substrate binding protein	Ion transport	1.5959	1.8220	1.6868	0.67	0.87	0.75
STN4L_00806	Peptide chain release factor 1	Protein biosynthesis		0.6642	0.5720		−0.59	−0.81
STN4L_00808	Adenosine deaminase	Purine metabolism	0.2546			−1.97		
STN4L_00813	A0A249DLE6 (uncharacterized protein)		2.1016	2.1606	2.5344	1.07	1.11	1.34
STN4L_00843	Cytoplasmic membrane protein			2.4288	2.1145		1.28	1.08
STN4L_00849	Type I restriction-modification system methyltransferase subunit	DNA metabolism	1.9960	1.5741		1.00	0.65	
STN4L_00892	Peptidase, U32 family	Proteolysis			0.4971			−1.01
STN4L_00915	Alkyl hydroperoxide reductase AhpD	Stress resistance	0.6054	0.4695	0.5354	−0.72	−1.09	−0.90
STN4L_00949	Formamidopyrimidine-DNA glycosylase	DNA metabolism (DNA repair)		2.2663			1.18	
STN4L_00956	Recombination protein RecR	DNA metabolism (DNA repair)	0.6130	0.5236	0.5763	−0.71	−0.93	−0.80
STN4L_00984	Branched-chain amino acid aminotransferase	Amino acid metabolism (BCAA)		1.8533	2.0413		0.89	1.03
STN4L_01000	3-Hydroxy-3-methylglutaryl coenzyme A reductase	Coenzyme metabolism	0.4641		0.5349	−1.11		−0.90
STN4L_01039	Sensor histidine kinase	Regulation process	1.8956	2.0681	1.7580	0.92	1.05	0.81
STN4L_01078	Peptidyl-prolyl *cis-trans* isomerase	Protein metabolism (folding)	1.5599		1.7045	0.64		0.77
STN4L_01188	Glycosyltransferase stabilizing protein Gtf2	Protein posttransductional modification			0.4793			−1.06
STN4L_01203	Restriction-modification enzyme type I M subunit	DNA metabolism			2.0223			1.02
STN4L_01253	Cysteine synthase	Amino acid metabolism (Cys)		0.6585	0.5902		−0.60	−0.76
STN4L_01260	Branched-chain amino acid ABC uptake transporter substrate-binding protein	Amino acid metabolism (BCAA)	1.5100	1.6713	1.5943	0.59	0.74	0.67
STN4L_01328	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Amino acid metabolism	2.2029	6.7933	1.8497	1.14	2.76	0.89
STN4L_01332	Nonribosomal peptide synthetase	Protein biosynthesis		0.5422	0.5889		−0.88	−0.76
STN4L_01366	Segregation and condensation protein B	Cell division	0.5634		0.6565	−0.83		−0.61
STN4L_01371	Noncanonical purine NTP pyrophosphatase	Purine degradation	0.5369	0.6540		−0.90	−0.61	
STN4L_01535	tRNA pseudouridine synthase A	tRNA processing			0.4070			−1.30
STN4L_01587	ATP-dependent Clp protease, ATP-binding subunit	Proteolysis			0.3778			−1.40
STN4L_01640	N^5-carboxyaminoimidazole ribonucleotide synthase	Purine metabolism	1.5012	1.9911	1.6300	0.59	0.99	0.70
STN4L_01642	Phosphoribosylamine-glycine ligase	Purine metabolism	2.1143	1.5531		1.08	0.64	

(Continued on next page)

TABLE 7 (Continued)

Locus	Annotation	Pathway	FC[a] (Y2/Y1)			log_2 FC		
			t3	t4	t5	t3	t4	t5
STN4L_01647	Bifunctional purine biosynthesis protein	Purine metabolism	2.3990	1.5389	1.5424	1.26	0.62	0.63
STN4L_01649	Phosphoribosylformylglycinamidine cyclo-ligase	Purine metabolism	2.5780	1.6917	2.0178	1.37	0.76	1.01
STN4L_01650	Amidophosphoribosyltransferase	Purine metabolism	2.6008	1.8975		1.38	0.92	
STN4L_01652	Phosphoribosylaminoimidazole-succinocarboxamide synthase	Purine metabolism	2.2791			1.19		
STN4L_01719	Tryptophan tRNA ligase	Protein biosynthesis			2.4853			1.31
STN4L_01735	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	tRNA processing	0.4810	0.6113		−1.06	−0.71	
STN4L_01777	50S ribosomal protein L33	Protein biosynthesis		0.5851	0.3693		−0.77	−1.44
STN4L_01827	30S ribosomal protein S3	Protein biosynthesis		0.5526	0.6529		−0.86	−0.62
STN4L_01830	30S ribosomal protein S17	Protein biosynthesis		0.6401	0.5758		−0.64	−0.80
STN4L_01831	50S ribosomal protein L14	Protein biosynthesis		0.6154	1.5496		−0.70	0.63
STN4L_01838	30S ribosomal protein S5	Protein biosynthesis		0.5839	0.5722		−0.78	−0.81
STN4L_01844	50S ribosomal protein L36	Protein biosynthesis	1.6367		1.5773	0.71		0.66
STN4L_01895	Acetolactate synthase	Amino acid metabolism (BCAA)	1.5508	1.5593		0.63	0.64	
STN4L_01963	Argininosuccinate synthase	Amino acid metabolism (Arg)		0.5898	0.3349		−0.76	−1.58
STN4L_02064	Phospho-2-dehydro-3-deoxyheptonate aldolase	Amino acid metabolism (Aro)			2.2735			1.18
STN4L_02079	Response regulator of the LytR/AlgR family protein	Regulation process	0.2526			−1.99		
STN4L_02080	Putative uncharacterized conserved secreted protein		0.3717	0.3812	0.5400	−1.43	−1.39	−0.89
STN4L_02142	C3-degrading proteinase	Proteolysis			0.3455			−1.53
STN4L_02202	ATP-dependent proteinase ATP-binding subunit	Proteolysis			0.1785			−2.49
STN4L_02204	A0A249DMF3 (phosphoribosylglycinamide synthetase)	Purine metabolism		1.6200	2.1677		0.70	1.12

[a]FC, fold change of Y2/Y1, restricted to the most significant changes (most intense and/or systematic over time growth).

TABLE 8 Main outcomes of the study

Category	Y1	Y2
Yeast extract composition		
Peptide		
Diversity	+	
Richness		+
Hydrophobicity	+	
Negative charge		+
Free amino acids		
Total amt	=	=
Consumption rate		+
BCAAs	=	=
Cellular response		
General pathways		
Amino acid supply		+ (R, P)[a]
Citrate metabolism	+ (R, P)	
Purine biosynthesis		+ (R, P)
Pyrimidine metabolism	+ (R)	
Ion transport		+ (P)
Stress resistance	+ (P)	
Specific metabolisms		
Urease metabolism		+ (R)
Streptide biosynthesis	+ (R)	
Regulation		
GntR (STN4L_00673)	+ (P)	
HK02 (STN4L_01039)		+ (P)
Transcriptional regulator STN4L_01175	+ (R)	
NdrR (STN4L_01309)	+ (P)	
LytR (STN4L_02079)	+ (R, P)	
CodY (STN4L_02180)		+ (R)

[a]R, RNA sequencing analysis; P, proteomics analysis.

duction processes involve a step of yeast protein hydrolysis that can be achieved by autolysis or addition of external proteases, and technical specificities such as physicochemical conditions of autolysis (temperature and pH) or specificities of the added proteases might affect the free amino acid and peptide contents of the YE. In most industrial culture media, YE is generally considered the main supply of nitrogen. Despite these differences between the two YEs in terms of amino acid and peptide concentrations (and compositions), similar growth was observed, indicating that both media equally fulfilled the nutritional requirements of the strain. Moreover, the higher rate of peptide consumption in Y2 did not result in a higher growth rate of the strain. This can be explained as follows: peptide transport rate depends on both the concentration of the peptide and its affinity for the oligopeptide transport system Ami. It has no relationship with the nutritional requirements of the strain. Therefore, the higher consumption rate in Y2 might be due to a higher peptide concentration. Moreover, the fact that the growth rate was not stimulated in Y2 compared to that in Y1 suggests that these transported peptides did not provide growth rate-limiting amino acids to the strain. This correlates with the suggestion made by Kevvai and coworkers that the incorporation of an amino acid in the peptide form by *L. lactis* is more dependent on the peptide availability than on the auxotrophy of the strain for the amino acid (4). Albeit the strain grew similarly in the two growth media, we observed significant variations in gene expression and protein synthesis profiles, corresponding to different responses of the strain to the two similar growth media (Table 8).

YE and CodY regulation. Amino acid biosynthesis pathways were differentially affected by the YE, with a general stimulation of gene expression and protein production in Y2 compared to that in Y1 at least in the exponential growth phase. At the end of growth, the situation was more contrasted, with stimulation of some pathways in Y1 (Gln transport and Arg biosynthesis) or in Y2 (His biosynthesis). In *L. lactis*, glutamine transport genes *glnPQ* are downregulated by the local regulator GlnR in response to an excess of glutamate in the growth medium, whereas argininosuccinate synthetase *argG*

is upregulated under the same condition (19). This does not correspond to what we observed (upregulation of all of these genes in Y1). So, despite the presence of a homolog of *glnR* in the N4L genome (*STN4L_02009*), the hypothesis that this differential gene expression between the two YEs was only due to differences in GlnR-dependent regulation appears unlikely.

The histidine biosynthesis pathway was more stimulated in Y2 than in Y1. The presence of histidine in the growth medium has been reported to play a role in the transcription of the *his* operon (20). Interestingly, His was present in the free form only in Y1, whereas it was only present in the peptide form in Y2 (Table 1). It therefore suggests that His-containing peptides present in Y2 could be less efficiently used as a source of histidine than free His, as biosynthesis of His was required in Y2 and not in Y1.

In addition to some targeted regulations, amino acid metabolism (biosynthesis and transport) in *Firmicutes* is known to be controlled by the master pleiotropic transcriptional regulator CodY (5). Transcriptional studies performed in *L. lactis* revealed, in particular, the repression of histidine biosynthesis-encoding genes by CodY (21, 22). Interestingly, *codY* (*STN4L_02180*) was detected as substantially differentially expressed. Because of the higher expression of *codY* in Y2, the genes under the control of CodY were expected to be expressed at a lower level in Y2 than in Y1. That was clearly the case for arginine and glutamine biosynthesis genes but not for histidine biosynthesis genes. The effect of CodY on regulated genes depends on its interaction with the so-called CodY box, which represents the CodY binding site. The CodY box has been identified in *S. thermophilus* ST2017 and contains 15 nucleotides [AA(T/A)(A/T)TTCTGA (A/C)AATT] (23). The CodY box of *S. thermophilus* N4L seems a bit degenerated (TTTTCCCTGAAAA). It was detected in the front of the *argG* gene (TTTTGACTAAAA), supporting the hypothesis of a repression of arginine biosynthesis in Y2 by CodY. In contrast, no CodY boxes were found in the promoter regions of His and Glu biosynthesis operons, therefore questioning their regulation processes in Y2.

Furthermore, genes involved in citrate metabolism were overexpressed in Y1. Citrate metabolism is reported to be repressed by CodY (21). Therefore, the differential expression/production of genes/proteins involved in citrate metabolism is in good agreement with the differential time course of CodY.

Is CodY the only pleiotropic regulator acting on nitrogen nutrition? The BCAA biosynthesis genes were reported to be the most tightly regulated by CodY in *S. thermophilus* (23). Surprisingly, no significant difference was evidenced in the expression levels of the BCAA biosynthesis-encoding genes (*ilvADBNCD1*, i.e., *STN4L_01489*, *-01894* to *-01897*, and *-02214*), which decreased during growth in Y1 and in Y2 to the same extent. Therefore, the similar time course of BCAA-encoding gene expression in the two YEs was not in agreement with the higher expression of *codY* in Y2. This apparent discrepancy, together with that revealed with histidine biosynthesis, strongly suggests that the CodY regulation cannot explain all the differences we observed during the growth in the two YE-based media.

The same study suggested that CodY does not control the expression of the *S. thermophilus* proteolytic system genes, namely, cell wall anchored proteinase, peptide transport systems, and intracellular peptidases (23). Under our experimental conditions, these genes were expressed at similar levels in the two YEs, regardless of the growth phase (see Table S5 in the supplemental material). As *codY* expression was higher in Y2 than in Y1, our results agreed with the reported absence of control of the proteolytic system-encoding genes by CodY. Nevertheless, the expression level of the proteolytic system-encoding genes decreased as a function of time in both media to the same extent. It therefore suggests regulation is mediated by a process other than with CodY.

YebC was recently described as a transcriptional repressor of key genes of the proteolytic system of *Lactobacillus* (24). As no homolog of CodY has been found in lactobacilli, YebC is regarded as the CodY-like transcriptional regulator responsible for the regulation of the proteolytic system gene expression in lactobacilli. The two transcriptional regulators are present in *S. thermophilus*, and YebC from *S. thermophilus*

shares 44% of homology with YebC from *Lactobacillus delbrueckii* subsp. *lactis*. Both streptococcal regulators YebC and CodY were not only expressed but also found as differentially expressed under our two growth conditions. It therefore questions their interplay during *S. thermophilus* growth as well as the reasons for these differential expressions.

YE and purine metabolism. The purine biosynthesis pathway was also significantly stimulated in Y2 compared to that in Y1, at both the gene and protein levels. This pathway is under the control of the local regulator PurR, whose repressive activity is stimulated by the presence of purines in the culture medium (25). Expression of genes involved in the *de novo* purine biosynthesis (*pur* operon) is required only in the presence of a high intracellular concentration of phosphoribosyl-pyrophosphate (PRPP), the precursor of IMP. In the absence of PRPP, PurR binds to promoter of the *pur* operon and inhibits the transcription initiation of *pur*-regulated genes (25). According to the model developed in *Bacillus subtilis*, addition of purine to the growth medium results in the intracellular accumulation of ADP that in turn results in the decrease of PRPP (26). As the production process of YEs is designed to enrich the purine content of Y2 (Procelys, personal communication), PurR regulation is expected to be more active in Y2 than in Y1. Here again, that is not what has been observed, and it does not sustain the hypothesis of a prominent regulative action of PuR on purine metabolism under our experimental conditions. Glutamine PRPP amidotransferase catalyzes the initial reaction in the *de novo* purine nucleotides synthesis, by transferring the amido nitrogen of glutamine to PRPP, yielding glutamate and 5-phosphoribosylamine. Thus, there is a link between nitrogen metabolism and nucleotide biosynthesis (27). Therefore, the differences between the Y1 and Y2 contents of free amino acids and peptide-bound amino acids, especially glutamine and glutamic acid, might also play a role in the differential expression of the *pur* operon genes. Accordingly, *de novo* purine nucleotide synthesis has been reported to be regulated, at least in part, by end product inhibition of glutamine PRPP amidotransferase, i.e., adenine and guanine (28). However, this statement suggests that purine nucleotide synthesis might also be regulated in a way other than by PurR. More recently, CodY was reported to induce expression of most genes of the purine biosynthesis pathway, including *purC*, *-D*, *-H*, *-F*, and *-M* (23). As discussed just above, *codY* was significantly more highly expressed during growth in Y2 than in Y1. Consequently, CodY could likely stimulate the expression of *pur* genes in Y2 to a larger extent than in Y1. This would suggest that CodY activation might override PurR repression and confirm that regulations of purine and nitrogen metabolisms are linked.

YE and urease. One of the functions of urease is to supply ammonia for the synthesis of glutamine (29). Arioli and coworkers demonstrated that glutamine and glutamate repressed expression of the *S. thermophilus* urease operon (30). As the amount of Glu in the peptide form was higher in Y2 than in Y1, whereas free Glu was more abundant in Y1 than in Y2, it therefore suggests that either the repressive effect of Glu in Y1 is mainly exerted by free Glu or that the Glu-containing peptides present in Y2 are less efficiently transported into the cells by the Ami system than their Y1 counterparts.

YE and quorum sensing. Another lesson from this global analysis was the detection of a significant number of expressed shCDS during growth of *S. thermophilus* in YE-containing media. One of the major known roles of such short coding sequences is to act as pheromone precursors inducing quorum sensing-controlled functions (7). A few examples of such functions have already been described in *S. thermophilus*: streptide production (10, 11), natural competence (31, 32), and bacteriocin production (33).

Streptide is a cyclic peptide whose production by some *S. thermophilus* strains is induced by a short hydrophobic peptide (SHP). Streptide production was largely stimulated in Y1 only. Indeed, the SHP pheromone-encoding shCDS (*STN4L_short_044*) was not expressed in Y2, whereas its expression gradually increased in Y1. Why did this induction take place during growth in Y1 and not in Y2? If we consider the scheme of

streptide production (a quorum sensing mechanism where induction of streptide gene expression occurs when the amount of the mature pheromone passes a threshold in the external medium and is imported back into the cell by the oligopeptide transporter Ami), the most probable explanation has to be found in the difference in peptide compositions between the two YEs. The biochemical properties (length, mean charge, and hydrophobicity) of peptides from Y1 and Y2 fit the preference of the Ami system of *S. thermophilus* (34). As Y2 contained a larger amount of peptides than Y1, it can be simply hypothesized that, in Y2, competition for peptide transport via the Ami system occurred so that the SHP pheromone could not be imported back and induce streptide production. Such a competition phenomenon for peptide transport responsible for quorum sensing disruption has already been described (35). Another hypothesis could be that some peptides from Y1 act as inducers of the quorum sensing system. SHP pheromones are octa- or nonamer peptides characterized by the presence of an acidic amino acid (Asp or Glu) at the N terminus of the mature sequence, a Gly residue at the C terminus, and a BCAA stretch of at least 3 residues within the mature sequence (36). Several SHP-like peptides that partly meet these features (in particular, presence of a stretch of BCAAs) have been identified in YEs (see Table S6). It cannot be excluded that at least one of these present in Y1 (and not in Y2) mimic the action of the SHP pheromone STN4L_short_044 and initiate streptide production. Actually, nonspecific nutritive oligopeptides derived from caseins are able to mimic the ComS pheromone action and to induce competence development in *S. thermophilus* (37). Finally, an opposite hypothesis can also be proposed. Y2 contained a large number of highly hydrophobic peptides, as indicated by the outliers in Fig. 1. It is therefore likely that one of these peptides present in Y2 (and not in Y1) acts as a quencher. That could be for instance the case of KPLDVVIPIG, a decamer peptide that possesses both a BCAA stretch and a Gly in the last position of the sequence but a positively charged residue (Lys) instead of a negatively charged amino acid at the N terminus. That could also be the case for DNEHLVLPR, a nonamer peptide that contains both a BCAA stretch and an acidic residue (Asp) in the first sequence position but a positively charged residue (Lys) instead of a Gly at the C terminus.

More generally, these hypotheses mean that YE-containing growth media could contain peptides that interfere with quorum sensing mechanisms, either by eliciting the quorum sensing response (pheromone-like peptides) or by quenching the quorum sensing response. To evaluate the relevance of this assumption, YE peptides have been screened for their possible regulatory roles, using the iQSp sequence-based tool dedicated to the prediction of quorum sensing peptides, based on their physicochemical properties (38). Half of the peptides (48% in Y1, 49% in Y2) were predicted to be potential quorum sensing peptides. This proportion was surprisingly high and most probably overestimated. However, it supports the idea that YEs might contain peptides able to induce quorum sensing pathways.

Concluding remarks. All the above taken together indicate that even when growing in similar growth media, *S. thermophilus* differentially expresses a number of genes, including transcriptional regulators. Some of the proteins encoded by these differentially expressed genes may have important biological or technological functions, such as stress resistance and nitrogen supply. It cannot be excluded that some of the differentially expressed shCDS also control some relevant physiological functions via quorum sensing mechanisms. Therefore, during the industrial process of starter propagation, the metabolic potential (defined as the pool of expressed genes during the propagation step) of the starter might be differentially programmed by two growth media that yield similar growth parameters. This differential metabolic programming might affect not only the behavior of the strain during further technological processes of starter production (e.g., cooling, concentration, freezing, and freeze-drying) but also some technological properties such as lag duration or acidification during subsequent growth in milk, depending on the genes differentially expressed.

MATERIALS AND METHODS

Strain and culture conditions. The proteinase-positive strain *S. thermophilus* N4L (NZ_LS974444.1) was provided by Sacco S.r.l. (Cadorago, Italy). It was routinely precultured in M17 broth (39) supplemented with 5% (wt/vol) lactose. Cultures were fermented in 1-liter bioreactors (BIOSTAT Qplus system; Sartorius Stedim Biotech, Germany). The fermenters were inoculated at 2% (vol/vol) with the following parameters: temperature of 40°C, agitation of 50 rpm, initial pH manually adjusted to 6.6 before inoculation then automatically regulated at 6.0 with 2 M sodium hydroxide, and nitrogen supply at 0.2 liters/min in the headspace. The culture medium contained (wt/vol) 6% lactose, 0.01% calcium chloride, 2% YE, and a pool of vitamins, as previously described (40). Growth was followed by optical density (600 nm) and by online monitoring of the volume of added NaOH. Two distinct YEs were used, both from the NuCel range (Procelys, Maisons-Alfort, France). They were named Y1 and Y2 throughout the text. Y1 and Y2 were produced from selected strains of *Saccharomyces cerevisiae* yeast, using specific industrial processes. In particular, Y2 displayed a higher nucleotide content than Y1 as per the manufacturer's settings (Procelys, personal communication).

Three replicate cultures were performed for each YE-based growth medium. Samples dedicated to transcriptomics and proteomics were collected after 3, 4, and 5 h of growth. To ensure maximal reproducibility, the exact moments of sampling for the 2nd and 3rd repetitions were adjusted according to the volumes of NaOH added so that they matched the sampling point of the 1st repetition. At each sampling point, the sample was divided into two aliquots dedicated to transcriptomics (6 ml) and proteomics (15 ml).

RNA sequencing and transcriptome analysis. After harvest, samples were rapidly aliquoted (3×2 ml), the supernatants were removed by fast centrifugation (18,000 $\times$ *g*, 15 s), and bacterial pellets were frozen in liquid nitrogen and stored at −80°C. Cells were lysed in phenol-chloroform 5:1 (vol/vol) with a FastPrep FT120 (Thermo Savant, USA). Two cycles of cell disruption were performed, both times at 6.5 m/s for 45 s. Cellular debris was pelleted by centrifugation (16,000 $\times$ *g*, 10 min, 4°C), and the RNA-containing aqueous supernatants were collected. Total RNA extraction was then performed using the Direct-Zol RNA miniprep kit (Zymo Research, USA) according to the manufacturer's instructions. Contaminant genomic DNA was removed with the DNA-free DNA removal kit (Invitrogen, USA). Purified RNAs were then subjected to an agarose gel-based migration in order to separate short transcripts (≈20 to 200 nt) from regular transcripts (>200 nt). The latter were depleted of rRNA using the Ribo-Zero rRNA removal kit (Epicentre, USA). Finally, assessment of RNA quantity and quality of both preparations (short and regular transcripts) was achieved using a Qubit fluorometer (Thermo Fisher, USA) and a Bioanalyzer system (Agilent, USA), respectively.

RNA sequencing was performed on the I2BC sequencing platform (http://www.i2bc.paris-saclay.fr). Two directional libraries were constructed from the purified samples. The short and regular transcript libraries were obtained using the TruSeqSmall RNA and the ScriptSeqRNA-seq library prep kits, respectively (Illumina, USA). Both libraries were then sequenced in single end (75-bp read length) on an Illumina NextSeq 500 platform. Raw reads were first demultiplexed and outputted to fastq files with bcl2fastq2 v. 2.15.0 (Illumina). They were subsequently trimmed using Cutadapt v. 1.9.1 (41), and their quality was controlled by FastQC v. 0.11.5. Cleaned reads were then mapped against *S. thermophilus* N4L genome (42) using BWA-aln v. 0.6.2 (43).

The N4L genome was processed beforehand using the BactgeneSHOW program to predict short putative coding sequences (shCDS) (16, 17). Options parameters were the following: -m 4C_si -rbs m1 -duprev -cdst 0.01. Predicted shCDS (45 to 180 nt) that were not present in the N4L published annotation were included in the data set (see Table S1 in the supplemental material). Following this work, relative expression of short and regular genes was computed with the featureCounts tool (option –Q 10) from the SubRead package v. 1.5.2 (44), using data from short and regular transcript-derived libraries, respectively (Table S2).

After this step, all abundance data were pooled and imported into the R SARTools environment v. 1.6.3 (45) in order to proceed to their normalization and differential analysis. For that purpose, DE-Seq2 v. 1.18.1 (46) was employed. Total read counts were first normalized by the median-of-ratios method (47). Gene expression in Y1 and Y2 media was then compared at each sampling time. The subsequently generated *P* values were adjusted by a false-discovery rate (FDR) procedure (48). Only genes showing an adjusted *P* value of ≤0.05 were considered statistically differentially expressed between Y1 and Y2 at the considered time point.

Proteome identification and quantification. Cellular fractions of collected samples were first pelleted by centrifugation (3,000 $\times$ *g*, 10 min, 4°C), washed twice in 200 mM $NaPO_4$ (pH 6.8) buffer, frozen in liquid nitrogen with a protease inhibitor cocktail (P8465; Sigma-Aldrich), and stored at −80°C until further use. Cells were disrupted with a Basic Z cell disruptor (CellD, Warwickshire, UK) at 2.5×10^5 kPa, and debris was removed by centrifugation (3,000 $\times$ *g*, 10 min, 4°C). The resulting supernatants were then ultracentrifuged (21,000 $\times$ *g*, 20 min, 4°C) to pellet an envelope protein-enriched fraction (EP). These pellets were separated from the cytosolic protein-containing supernatants (CP) and resuspended in phosphate buffer containing the protease inhibitor cocktail. Proteins in both CP and EP samples were separated by SDS-PAGE and digested in-gel by trypsin as previously described (49).

Tryptic peptides were analyzed at the PAPPSO platform (http://pappso.inra.fr) by LC-MS/MS using an UltiMate 3000 RSLCnano System (Thermo Fisher) coupled to an LTQ-Orbitrap Discovery mass spectrometer (Thermo Fisher). Five microliters of the tryptic peptide extracts was separated on a PepMap 100 C_{18} column (500 mm by 75 μm, 3 μm, 100 Å; Thermo Fisher) using a 2-slope gradient of acetonitrile (ACN) in 0.1% formic acid (from 3% to 19% ACN in 25 min and then to 80% in 10 min) at a 20-μl/min flow rate. Eluted peptides were then ionized with a 1.3 kV spray voltage applied to an uncoated capillary probe

(10-μm internal diameter; New Objective, USA). Precursor masses were first analyzed at high resolution (*m/z* 300 to 1,400, resolution = 15,000) in the Orbitrap analyzer. The eight most intense doubly and triply charged ions were subsequently fragmented in the LTQ linear trap via collision-induced dissociation (CID; collision energy of 35%). An exclusion window of 30 s was used to reduce identification redundancy.

Peptide identification and parental protein inference were finally performed with X!Tandem search engine v. 2015.12.15.2 (Vengeance) enclosed in the open-source software X!TandemPipeline v. 3.4.3 (50). EP and CP samples were analyzed separately (two final data sets). The peptide search was performed on the protein sequence of *S. thermophilus* N4L with the following constraints: tryptic cleavage specificity, cysteine carbamidomethylation, variable methionine oxidation state, and mass tolerance of $\pm$10 ppm and $\pm$0.5 Da for parental and fragment ions, respectively. Only one missed cleavage site was allowed. A minimum of four fragment ions was required for a peptide to be scored. A supplementary refinement step was also performed for a more in-depth search on peptides showing an E value of $\leq$0.01 to assess variable acetylation of peptide N-terminal residues. In the end, all peptide matches showing an E value of $\leq$0.05 were conserved, and only parental proteins identified with two peptides or more were retained. The false-discovery rate was evaluated using the reversed protein database of the N4L strain.

Identified peptides were quantified by the extracted-ion chromatography (XIC) method. MassChroq v. 2.2.1 (51) was used to align chromatograms, remove artefactual spikes, and detect and quantify peaks. Raw peptide intensities obtained were then imported in the R package MassChroqR v. 0.3.8 and $\log_{10}$ transformed for subsequent analysis according to the pipeline (peak filtration on the basis of retention times features and normalization of the remaining peaks). Protein-specific peptides detected in all samples (no missing data) were retained for protein relative quantification. Only proteins associated with at least two peptides were quantified by summing the XIC intensities of their constitutive peptides. The subcellular localization of identified proteins was investigated using the tool LocateP v.2 (52).

In the same way as for transcriptome analysis, protein relative abundance was compared between bacterial samples obtained from Y1 and Y2 at the same sampling point. An analysis of variance (ANOVA) was performed on proteins showing more than 50% of variation between compared conditions. Obtained *P* values were adjusted by an FDR procedure, and a 5% cutoff threshold was chosen to detect significant variations.

Functional annotation, overrepresentation test, and pathway representation. The genome of *S. thermophilus* N4L was functionally annotated using the KEGG Orthology and Links Annotation (Blast-KOALA) tool (53). Protein sequences of *S. thermophilus* N4L genes were blasted against the KEGG orthology (KO) database at the *Streptococcus* level, and hits were attributed a unique KO number.

Overrepresentation tests were carried out using the enrichKEGG function from the R package clusterProfiler v. 3.10.1 (54). Briefly, a hypergeometric test was used in order to determine whether the frequency of a specific term in the tested data set was significantly higher than that of a background distribution. The generated *P* values were corrected by FDR, and a molecular function was considered significantly overrepresented at a threshold of 5%.

Kinetic analysis of free and peptide-bound amino acids in the growth medium. A kinetic analysis of free and peptide-bound amino acid content of the growth media was performed during a single fermentation of *S. thermophilus* N4L in Y1 and Y2. Samples were collected every hour, centrifuged (3,000 $\times$ *g*, 10 min, 4°C), and filtered with a 0.22-μm-pore-sized polyvinylidene difluoride (PVDF) membrane with low protein binding properties (Millipore, USA). For free amino acid measurements, proteins and peptides present in the YE growth medium were precipitated in 3% (vol/vol) sulfosalicylic acid. Samples were stored overnight at 4°C and then centrifuged at 10,000 rpm for 10 min. Supernatants were diluted (1/50 [vol/vol]) in sample buffer containing 0.2 M boric acid and 0.2 M NaOH (1/1 [vol/vol]). For peptide-bound amino acid measurement, proteins present in the medium were precipitated by 1% trifluoroacetic acid (TFA) (vol/vol). Peptide content was then determined indirectly by quantifying the total amino acid content of the samples after acidic hydrolysis and subtracting the free amino acid content. A volume of 100 μl of each supernatant was dried overnight in a SpeedVac system. Dried extracts were hydrolyzed at 110°C for 16 h in the presence of 6 N HCl and dried again for 2 h to remove traces of acid. They were finally resuspended in 1 ml of neutralizing solution (100 g/liter NaOH, 74 g/liter KCl, and 61 g/liter boric acid) and diluted (1/5) in sample buffer.

Amino acid quantification relied on a precolumn derivatization with *o*-phthalaldehyde (OPA). Derivatized samples were loaded on a C_{18} Core-Shell Kinetex column (150 by 4.6 mm, 2.6 μm; Phenomenex, USA). The separation gradient was as follows: 0 to 10 min, 100% A; 10 to 18 min, 0% to 35% B; 18 to 28 min, 35% to 38% B; 28 to 50 min, 38% to 60% B; 50 to 55 min, 60% to 100% B; 60 to 65 min, 100% B to 100% A; 65 to 80 min, 100% A (solvent A: 3% [vol/vol] tetrahydrofurane [THF], 50 mM sodium acetate [pH 6.4]; solvent B: 5% [vol/vol] THF in methanol). Derivatized amino acids were detected by UV measurement at 337 nm.

Growth medium peptide identification. Oligopeptides (longer than 6 residues, due to technical constraints) initially present in Y1 and Y2 were identified as previously described (2). Briefly, 0.1% TFA (vol/vol) and 5% ACN (vol/vol) were added to the peptide-containing culture medium which was successively ultrafiltered through 10-kDa- and 3-kDa-cutoff Amicon Ultracel membranes (Millipore). Peptides were then extracted by solid-phase extraction using StrataX cartridges (Phenomenex) with an elution buffer consisting of 50% ACN (vol/vol) and 0.1% TFA (vol/vol). Eluted oligopeptides were dried, resuspended in 0.1% TFA (vol/vol), and ultrafiltered once again through a 3-kDa-pore-size Ultracel membrane (Millipore).

Oligopeptides were consecutively submitted to a 2D-HPLC separation procedure. The amount of peptides analyzed corresponded to 5 μg of YE. Peptides were first loaded on a Nucleoshell RP 18plus RP-HPLC column (150 by 4.6 mm, 2.7 μm; Macherey-Nagel, Germany) and eluted at 40°C with a linear

gradient of 1.6% ACN per min in ammonium formate (20 mM, pH 6.2). The flow rate was 0.7 ml/min. Peptides eluting between 5 and 25 min were collected in 0.7-ml-containing fractions. Each fraction was dried overnight, resuspended in 0.1% TFA and 2% ACN, and loaded onto a PepMap 100 C_{18} column (Thermo Fisher). Peptides of each fraction, corresponding to a total of 0.8 μg of YE, were finally separated and identified as described above, except that no cleavage specificity instead of tryptic cleavage was defined for peptide identification.

Bioinformatic tools. To evaluate if some of the YE peptides could act as a pheromone or a quorum quencher, we used the iQSp tool developed by Charoenkwan and coworkers (38). The CodY box motif was identified using the MEME software from the MEME suite (55) and promoter sequences of *codY* and *livJ* genes as a template (http://meme-suite.org/). Detection of CodY box in promoter sequences was performed using the FIMO software from the MEME-suite.

Data availability. The RNA sequencing raw data from this study have been deposited in the ArrayExpress database at EMBL-EBI under accession number E-MTAB-9434 (www.ebi.ac.uk/arrayexpress). The mass spectrometry proteomics and peptidomics data are also freely available from the PAPPSO PROTICdb database (56) with the project names N4L_proteome (http://moulon.inra.fr/protic/n4l_proteome) and YE_peptides (http://moulon.inra.fr/protic/ye_peptides).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.8 MB.

SUPPLEMENTAL FILE 4, XLSX file, 0.4 MB.

SUPPLEMENTAL FILE 5, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 6, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

We thank the INRAE PAPPSO proteomics platform (http://pappso.inrae.fr) for providing mass spectrometry facilities, the INRAE MIGALE bioinformatics platform (http://migale.inrae.fr) for providing computational resources, and the sequencing platform of the I2BC institute (https://www.i2bc.paris-saclay.fr) for conducting RNA sequencing experiments and for their help regarding RNA sequencing data analysis. We also thank C. Delorme, R. Gardan, and F. Rul (INRAE, MICALIS) for fruitful discussions.

L.P. performed the experimental study; E.H. performed the mass spectrometry experiments. L.P. and V.J. wrote the manuscript. All authors contributed to conception and design of the study, to manuscript revision, and approved the submitted version.

This work was funded by the Association Nationale de la Recherche et de la Technologie (ANRT, contract Nr 2015/0599).

Alain Sourabié and Iris Besançon are employed by Procelys; Martin Pedersen is employed by Sacco S.r.l. All other authors declare no competing interests.

REFERENCES

1. Takalloo Z, Nikkhah M, Nemati R, Jalilian N, Sajedi RH. 2020. Autolysis, plasmolysis and enzymatic hydrolysis of baker's yeast (*Saccharomyces cerevisiae*): a comparative study. World J Microbiol Biotechnol 36:68. https://doi.org/10.1007/s11274-020-02840-3.
2. Proust L, Sourabié A, Pedersen M, Besançon I, Haudebourg E, Monnet V, Juillard V. 2019. Insights into the complexity of yeast extract peptides and their utilization by *Streptococcus thermophilus*. Front Microbiol 10:906. https://doi.org/10.3389/fmicb.2019.00906.
3. Selby Smith J, Hillier AJ, Lees GJ, Jago GR. 1975. The nature of the stimulation of the growth of *Streptococcus lactis* by yeast extract. J Dairy Res 42:123–138. https://doi.org/10.1017/s0022029900015156.
4. Kevvai K, Kütt M-L, Nisamedtinov I, Paalme T. 2014. Utilization of ^{15}N-labelled yeast hydrolysate in *Lactococcus lactis* IL1403 culture indicates co-consumption of peptide-bound and free amino acids with simultaneous efflux of free amino acids. Antonie Van Leeuwenhoek 105:511–522. https://doi.org/10.1007/s10482-013-0103-2.
5. Guédon E, Serror P, Ehrlich SD, Renault P, Delorme C. 2001. Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in *Lactococcus lactis*. Mol Microbiol 40:1227–1239. https://doi.org/10.1046/j.1365-2958.2001.02470.x.
6. Shivers RP, Sonenshein AL. 2004. Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. Mol Microbiol 53:599–611. https://doi.org/10.1111/j.1365-2958.2004.04135.x.
7. Monnet V, Juillard V, Gardan R. 2016. Peptide conversation in Gram-positive bacteria. Crit Rev Microbiol 42:339–351. https://doi.org/10.3109/1040841X.2014.948804.
8. Perez-Pascual D, Monnet V, Gardan R. 2016. Bacterial cell-cell communication in the host via RRNPP peptide-binding regulators. Front Microbiol 7:706. https://doi.org/10.3389/fmicb.2016.00706.
9. Neiditch MB, Capodagli GC, Prehna G, Federle M. 2017. Genetic and structural analyses of RRNPP intercellular peptide signaling of Gram-positive bacteria. Annu Rev Genet 51:311–333. https://doi.org/10.1146/annurev-genet-120116-023507.
10. Ibrahim M, Guillot A, Wessner F, Algaron F, Besset C, Courtin P, Gardan R, Monnet V. 2007. Control of the transcription of a short gene encoding a cyclic peptide in *Streptococcus thermophilus*: a new quorum-sensing system? J Bacteriol 189:8844–8854. https://doi.org/10.1128/JB.01057-07.
11. Fleuchot B, Gitton C, Guillot A, Vidic J, Nicolas P, Besset C, Fontaine L, Hols P, Leblond-Bourget N, Monnet V, Gardan R. 2011. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. Mol Microbiol 80:1102–1119. https://doi.org/10.1111/j.1365-2958.2011.07633.x.
12. Schramma KR, Bushin LB, Seyedsayamdost MR. 2015. Structure and

biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. Nat Chem 7:431–437. https://doi.org/10.1038/nchem.2237.
13. Potvin J, Fonchy E, Conway J, Champagne CP. 1997. An automatic turbidimetric method to screen yeast extracts as fermentation nutrient ingredients. J Microbiol Methods 29:153–160. https://doi.org/10.1016/S0167-7012(97)00032-8.
14. Ummadi M, Curic-Bawden M. 2010. Use of protein hydrolysates in industrial starter culture fermentations, p 91–114. *In* Pasupuleti VK, Demain AL (ed), Protein hydrolysates in biotechnology. Springer, Dordrecht, Netherlands. https://doi.org/10.1007/978-1-4020-6674-0_6.
15. Khakimov B, Christiansen LD, Heins A-L, Sorensen KM, Schöller C, Clausen A, Skov T, Gernaey KV, Engelsen SB. 2017. Untargeted GC-MS metabolomics reveals changes in the metabolite dynamics of industrial scale batch fermentations of *Streptococcus thermophilus* broth. Biotechnol J 12:1700400. https://doi.org/10.1002/biot.201700400.
16. Nicolas P, Bize L, Muri F, Hoebeke M, Rodolphe F, Ehrlich SD, Prum B, Bessières P. 2002. Mining *Bacillus subtilis* chromosome heterogeneities using hidden Markov models. Nucleic Acids Res 30:1418–1426. https://doi.org/10.1093/nar/30.6.1418.
17. Ibrahim M, Nicolas P, Bessières P, Bolotin A, Monnet V, Gardan R. 2007. A genome-wide survey of short coding sequences in streptococci. Microbiology 153:3631–3644. https://doi.org/10.1099/mic.0.2007/006205-0.
18. Thévenard B, Rasoava N, Fourcassié P, Monnet V, Boyaval P, Rul F. 2011. Characterization of *Streptococcus thermophilus* two-components systems: *in silico* analysis, functional analysis and expression of response regulator genes in pure or mixed culture with its yogurt partner, *Lactobacillus delbrueckii* subsp. *bulgaricus*. Int J Food Microbiol 151:171–181. https://doi.org/10.1016/j.ijfoodmicro.2011.08.019.
19. Larsen R, Kloosterman TG, Kok J, Kuipers OP. 2006. GlnR-mediated regulation of nitrogen metabolism in *Lactococcus lactis*. J Bacteriol 188:4978–4982. https://doi.org/10.1128/JB.00025-06.
20. Delorme C, Ehrlich SD, Renault P. 1999. Regulation of expression of the *Lactococcus lactis* histidine operon. J Bacteriol 181:2026–2037. https://doi.org/10.1128/JB.181.7.2026-2037.1999.
21. den Hengst CD, van Hijum SAFT, Geurts JMW, Nauta A, Kok J, Kuipers OP. 2005. The *Lactococcus lactis* CodY regulon. Identification of a conserved *cis*-regulatory element. J Biol Chem 280:34332–34342. https://doi.org/10.1074/jbc.M502349200.
22. Guédon E, Sperandio B, Pons N, Ehrlich SD, Renault P. 2005. Overall control of nitrogen metabolism in *Lactococcus lactis* by CodY, and possible models for CodY regulation in *Firmicutes*. Microbiology 151:3895–3909. https://doi.org/10.1099/mic.0.28186-0.
23. Lu WW, Wang Y, Wang T, Kong J. 2015. The global regulator CodY in *Streptococcus thermophilus* controls the metabolic network for escalating growth in the milk environment. Appl Environ Microbiol 81:2349–2358. https://doi.org/10.1128/AEM.03361-14.
24. Brown L, Villegas JM, Elean M, Fadda S, Mozzi F, Saavedra L, Hebert EM. 2017. YebC, a putative transcriptional factor involved in the regulation of the proteolytic system of *Lactobacillus*. Sci Rep 7:8579. https://doi.org/10.1038/s41598-017-09124-1.
25. Kilstrup M, Hammer K, Ruhdal Jensen P, Martinussen J. 2005. Nucleotide metabolism and its control in lactic acid bacteria. FEMS Microbiol Rev 29:555–590. https://doi.org/10.1016/j.femsre.2005.04.006.
26. Weng M, Nagy PL, Zalkin H. 1995. Identification of the *Bacillus subtilis pur* operon repressor. Proc Natl Acad Sci U S A 92:7455–7459. https://doi.org/10.1073/pnas.92.16.7455.
27. Massière F, Badet-Denisot M-A. 1998. The mechanism of glutamine-dependent amidotransferases. Cell Mol Life Sci 54:205–222. https://doi.org/10.1007/s000180050145.
28. Chen S, Tomchick DR, Wolle D, Hu P, Smith JL, Switzer RL, Zalkin H. 1997. Mechanism of the synergistic end-product regulation of bacillus subtilis glutamine phosphorybosylpyrophosphatase amidotransferase by nucleotides. Biochemistry 36:10718–10726. https://doi.org/10.1021/bi9711893.
29. Monnet C, Mora D, Corrieu G. 2005. Glutamine synthesis is essential for growth of *Streptococcus thermophilus* in milk and is linked to urea catabolism. Appl Environ Microbiol 71:3376–3378. https://doi.org/10.1128/AEM.71.6.3376-3378.2005.
30. Arioli S, Monnet C, Guglielmetti S, Parini C, De Noni I, Hogenboom J, Halami PM, Mora D. 2007. Aspartate biosynthesis is essential for the growth of *Streptococcus thermophilus* in milk, and aspartate availability modulates the level of urease activity. Appl Environ Microbiol 73:5789–5796. https://doi.org/10.1128/AEM.00533-07.
31. Gardan R, Besset C, Guillot A, Gitton C, Monnet V. 2009. The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. J Bacteriol 191:4647–4655. https://doi.org/10.1128/JB.00257-09.
32. Fontaine L, Boutry C, Henry de Frahan M, Delplace B, Fremaux C, Horvath P, Boyaval P, Hols P. 2010. A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. J Bacteriol 192:1444–1454. https://doi.org/10.1128/JB.01251-09.
33. Fontaine L, Boutry C, Guédon E, Guillot A, Ibrahim M, Grossiord B, Hols P. 2007. Quorum-sensing regulation of the production of Blp bacteriocins in *Streptococcus thermophilus*. J Bacteriol 189:7195–7205. https://doi.org/10.1128/JB.00966-07.
34. Juille O, Le Bars D, Juillard V. 2005. The specificity of oligopeptide transport by *Streptococcus thermophilus* resembles that of *Lactococcus lactis* and not that of pathogenic streptococci. Microbiology 151:1987–1994. https://doi.org/10.1099/mic.0.27730-0.
35. Gardan R, Besset C, Gitton C, Guillot A, Fontaine L, Hols P, Monnet V. 2013. Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. J Bacteriol 195:1845–1855. https://doi.org/10.1128/JB.02196-12.
36. Fleuchot B, Guillot A, Mézange C, Besset C, Chambellon E, Monnet V, Gardan R. 2013. Rgg-associated SHP signaling peptides mediate cross-talk in streptococci. PLoS One 8:e66042. https://doi.org/10.1371/journal.pone.0066042.
37. Fontaine L, Goffin P, Dubout H, Delplace B, Baulard A, Lecat-Guillet N, Chambellon E, Gardan R, Hols P. 2013. Mechanism of competence activation by the ComRS signalling system in streptococci. Mol Microbiol 87:1113–1132. https://doi.org/10.1111/mmi.12157.
38. Charoenkwan P, Schaduangrat N, Nantasenamat C, Piacham T, Shoombuatong W. 2019. iQSP, a sequence-based tool for the prediction and analysis of quorum sensing peptides via Chou's 5-steps rule and informative physicochemical properties. Int J Mol Sci 21:75. https://doi.org/10.3390/ijms21010075.
39. Terzaghi BE, Sandine WE. 1975. Improved medium for lactic streptococci and their bacteriophages. Appl Microbiol 29:807–813. https://doi.org/10.1128/AEM.29.6.807-813.1975.
40. Letort C, Juillard V. 2001. Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus*. J Appl Microbiol 91:1023–1029. https://doi.org/10.1046/j.1365-2672.2001.01469.x.
41. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
42. Proust L, Loux V, Martin V, Magnabosco C, Pedersen M, Monnet V, Juillard V. 2018. Complete genome sequence of the industrial fast-acidifying strain *Streptococcus thermophilus* N4L. Microbiol Resour Announc 7:e01029-18. https://doi.org/10.1128/MRA.01029-18.
43. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
44. Liao Y, Smyth GK, Shi W. 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 41:e108. https://doi.org/10.1093/nar/gkt214.
45. Varet H, Brillet-Guéguen L, Coppée J-Y, Dillies M-A. 2016. SARTools: a DESeq2- and EdgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. PLoS One 11:e0157022. https://doi.org/10.1371/journal.pone.0157022.
46. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8.
47. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome Biol 11:R106. https://doi.org/10.1186/gb-2010-11-10-r106.
48. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
49. Henry C, Haller L, Blein-Nicolas M, Zivy M, Canette A, Verbrugghe M, Mézange C, Boulay M, Gardan R, Samson S, Martin V, André-Leroux G, Monnet V. 2019. Identification of Hanks-type kinase PknB-specific targets in the *Streptococcus thermophilus* phosphoproteome. Front Microbiol 10:1329. https://doi.org/10.3389/fmicb.2019.01329.
50. Langella O, Valot B, Balliau T, Blein-Nicolas M, Bonhomme L, Zivy M. 2017. XTandemPipeline: a tool to manage sequence redundancy for protein inference and phosphosite identification. J Proteome Res 16:494–503. https://doi.org/10.1021/acs.jproteome.6b00632.

51. Valot B, Langella O, Nano E, Zivy M. 2011. MassChroQ: a versatile tool for mass spectrometry quantification. Proteomics 11:3572–3577. https://doi.org/10.1002/pmic.201100120.
52. Zhou M, Boekhorst J, Francke C, Siezen RJ. 2008. LocateP: genome-scale subcellular-location predictor for bacterial proteins. BMC Bioinformatics 9:173. https://doi.org/10.1186/1471-2105-9-173.
53. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006.
54. Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16:284–287. https://doi.org/10.1089/omi.2011.0118.
55. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME suite: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208. https://doi.org/10.1093/nar/gkp335.
56. Langella O, Zivy M, Joets J. 2007. The PROTICdb database for 2-DE proteomics. Methods Mol Biol 355:279–303. https://doi.org/10.1385/1-59745-227-0:279.